NO-ARBITRAGE BOUNDS FOR THE FORWARD SMILE GIVEN MARGINALS

SERGEY BADIKOV, ANTOINE JACQUIER, DAPHNE QING LIU AND PATRICK ROOME

Abstract. We explore the robust replication of forward-start straddles given quoted (Call and Put options) market data. One approach to this problem classically follows semi-infinite linear programming arguments, and we propose a discretisation scheme to reduce its dimensionality and hence its complexity. Alternatively, one can consider the dual problem, consisting in finding optimal martingale measures under which the upper and the lower bounds are attained. Semi-analytical solutions to this dual problem were proposed by Hobson and Klimmek [14] and by Hobson and Neuberger [15]. We recast this dual approach as a finite dimensional linear programme, and reconcile numerically, in the Black-Scholes and in the Heston model, the two approaches.

1. Introduction

Since David Hobson's seminal contribution [12], an important stream of literature has focused on developing model-free sub(super)-hedges for multi-dimensional derivative products or path-dependent options, given a set of European option instruments. The key observation is that the model-free sub(super)-hedging cost is closely related to the Skorokhod Embedding problem (see the exhaustive survey papers by Hobson [13] and Obłój [18] in the context of mathematical finance). Recently, Beiglböck, Henry-Labordère and Penkner [3] studied this problem in the framework of martingale optimal transport theory. Assuming that European Call/Put option prices are known for all strikes and some maturities (equivalently the marginal distributions of the asset price are known at these times), optimal transport then yields a no-arbitrage range of prices of a derivative product consistent with these marginal distributions. The primal problem endeavours to find the supremum, or the infimum, of these prices over all joint martingale measures (transport plans) consistent with the marginals. The dual problem, in turn, seeks to find the 'best' sub(super)-replicating portfolio; this dual formulation has the advantage of a natural financial interpretation and can be cast as an (infinite) linear programme, amenable to numerical implementations, as proposed in [10].

Forward-start options (of Type I and of Type II) are among the simplest products amenable to these techniques. The upper bound price for the at-the-money Type-II forward-start straddle (with payoff $|S_{t+\tau} - S_t|$ for some $t, \tau > 0$) was computed by Hobson and Neuberger [15], where the support of the optimal martingale measure is a binomial tree. Unfortunately the optimal measure and the associated super-hedging portfolio are not available analytically. The martingale optimal transference plan for the lower bound price of the at-the-money Type-II forward-start straddle has been characterised semi-analytically by Hobson and Klimmek [14], and the transference plan (supported along a trinomial tree) is found by solving a set of coupled ODEs. Recently, Campi, Laachir and Martini [7] studied the change of numeraire in these two-dimensional optimal transport problems

Date: November 26, 2016.

2010 *Mathematics Subject Classification.* 91G20, 91G80, 90C46, 90C05, 90C34.

Key words and phrases. martingale optimal transport, robust bounds, forward-start, heston.

The authors are indebted to Claude Martini for stimulating discussions, and to the anonymous referees for helpful suggestions. AJ acknowledges financial support from the EPSRC First Grant EP/M008436/1.

and showed that, under some technical conditions, the lower bound for the Type-I at-the-money forward-start straddle is also attained by the Hobson-Klimmek transference plan.

In this paper, we numerically investigate the no-arbitrage bounds of the Type-II forward-start straddle. The infinite-dimensional linear programme corresponding to the optimal transport problem is presented in Section 2. Section 3 focuses on a reduction of the dimension of the problem, by discretising the support of the marginal distributions at times t and $t+\tau$, which respects the consistency of the primal and the dual problems with observed option prices, and yields robust numerical results. Our discretisation method differs from that of Henry-Labordère [10], and requires far fewer points, thus reducing the complexity of the finite-dimensional linear programmes to be solved, thereby improving the algorithmic speed. In Section 5, we specialise our computation of the upper and lower bounds to the cases where the marginal distributions are generated from a Black-Scholes model (lognormal marginals) and from the Heston stochastic volatility model. In the lower bound at-the-money case we numerically solve, in Section 4, the coupled ordinary differential equations associated with the Hobson-Klimmek transference plan, and show that it is in striking agreement with the LP solution of the dual problem. In Section 6 we numerically solve the primal problem and provide the optimal transport plans for a large range of strikes. The transport plans are only known for the at-the-money case [14, 15], and we highlight numerical evidence that the optimal transference plans are more subtle in these cases, and appear to be a combination of the lower and upper bound at-the-money plans. Intuitively, the extremal measure yields a price corresponding either to the maximum or to the minimum value of the product. In the forward-start option case, this extremal measure maximises or minimises the kurtosis of the conditional distribution of the asset price process (see in particular Sections 4 and 6 and [14, 15]). Therefore a choice of a model that misspecifies the kurtosis might lead to wrong risk exposure profile.

In the examples explored in Section 5 the range of forward smiles consistent with the marginal laws is large, and of the same magnitude as in [10, Section 5.5]. This wide range of prices supports the claim that using European vanilla options to replicate forward volatility-dependent claims seems illusory. Forward-start options should therefore be seen as fundamental building blocks for exotic option pricing and not decomposable (or approximately decomposable) into European options. Forward-start options are usually available from the stripping of cliquet options, on OTC markets. Cliquet options have been a popular instrument on Equity markets [6], and also as a hedging tool for insurance companies (the so-called FLEX Index options, the details of which can be found on the CBOE website). Models used for forward volatility-dependent exotics should be able to calibrate forward-start option prices and should produce realistic forward smiles consistent with trader expectations and observable prices.

Notations: $B_b(\mathbb{R})$ denotes the set of bounded measurable functions on the real line, $\mathbb{R}_+ := [0,\infty)$ represents the positive half-line; $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product, and $\|\cdot\|$ the L^1 norm.

2. Problem formulation

We consider an asset price process $(S_t)_{t\geq 0}$ starting at $S_0 = 1$, and we assume that interest rates and dividends are null. For $t,\tau > 0$, let μ and ν denote the distributions of S_t and $S_{t+\tau}$, assumed to have common finite mean, supported on $[0,\infty)$, and absolutely continuous with respect to the Lebesgue measure. We say that the bivariate law ζ is a martingale coupling, and write $\zeta \in \mathcal{M}(\mu,\nu)$, if ζ has marginals μ and ν and $\int_{y\in\mathbb{R}_+}(y-x)\zeta(\mathrm{d}x,\mathrm{d}y) = 0$ for each $x \in \mathbb{R}_+$. Following [22], we shall say that μ and ν are in convex, or balayage, order (which we denote $\mu \preceq \nu$) if they have equal means and satisfy $\int_{\mathbb{R}_+}(y-x)^+\mu(\mathrm{d}y) \leq \int_{\mathbb{R}_+}(y-x)^+\nu(\mathrm{d}y)$ for all $x \in \mathbb{R}$. This assumption

ensures (see [2]) that the set $\mathcal{M}(\mu,\nu)$ is not empty. Our objective is to find the tightest possible lower and upper bounds, consistent with the two marginal distributions, for the forward-start straddle payoff $|S_{t+\tau} - \mathfrak{K}S_t|$ with $\mathfrak{K} > 0$. To this end we define our primal problems:

$$\underline{\mathcal{P}}(\mu,\nu) := \inf_{\zeta\in\mathcal{M}(\mu,\nu)} \int_{\mathbb{R}_+^2} |y - \mathfrak{K}x|\zeta(\mathrm{d}x,\mathrm{d}y), \qquad \text{and} \qquad \overline{\mathcal{P}}(\mu,\nu) := \sup_{\zeta\in\mathcal{M}(\mu,\nu)} \int_{\mathbb{R}_+^2} |y - \mathfrak{K}x|\zeta(\mathrm{d}x,\mathrm{d}y). \tag{2.1}$$

We now define the following sub- and super-replicating portfolios:

$$\underline{Q} := \left\{(\psi_0,\psi_1,\delta) \in L^1(\mu)\times L^1(\nu)\times B_b(\mathbb{R}_+) : \psi_1(y) + \psi_0(x) + \delta(x)(y-x) \le |y-\mathfrak{K}x|, \text{for all } x,y\in\mathbb{R}_+\right\},$$
$$\overline{Q} := \left\{(\psi_0,\psi_1,\delta) \in L^1(\mu)\times L^1(\nu)\times B_b(\mathbb{R}_+) : \psi_1(y) + \psi_0(x) + \delta(x)(y-x) \ge |y-\mathfrak{K}x|, \text{for all } x,y\in\mathbb{R}_+\right\}.$$

Clearly if $(\psi_0,\psi_1,\delta) \in \underline{Q}$ (respectively $\in \overline{Q}$) then $\int_{\mathbb{R}_+^2} |y-\mathfrak{K}x|\zeta(\mathrm{d}x,\mathrm{d}y) \ge (\le) \int_{\mathbb{R}_+}\psi_0(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+}\psi_1(y)\nu(\mathrm{d}y)$. The dual problems are then defined as the supremum (infimum) over all sub(super)-replicating portfolios:

$$\begin{aligned}
\int_{\mathbb{R}_+^2} |y-\mathfrak{K}x|\zeta(\mathrm{d}x,\mathrm{d}y) &\ge \sup_{(\psi_0,\psi_1,\delta)\in\underline{Q}} \left\{\int_{\mathbb{R}_+}\psi_0(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+}\psi_1(y)\nu(\mathrm{d}y)\right\} =: \underline{\mathcal{D}}(\mu,\nu),\\
\int_{\mathbb{R}_+^2} |y-\mathfrak{K}x|\zeta(\mathrm{d}x,\mathrm{d}y) &\le \inf_{(\psi_0,\psi_1,\delta)\in\overline{Q}} \left\{\int_{\mathbb{R}_+}\psi_0(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+}\psi_1(y)\nu(\mathrm{d}y)\right\} =: \overline{\mathcal{D}}(\mu,\nu).
\end{aligned} \tag{2.2}$$

In [3, Theorem 1 and Corollary 1.1], the authors proved (actually for a more general class of payoff functions) that there is no duality gap, namely that both equalities $\underline{\mathcal{P}}(\mu,\nu) = \underline{\mathcal{D}}(\mu,\nu)$ and $\overline{\mathcal{P}}(\mu,\nu) = \overline{\mathcal{D}}(\mu,\nu)$ hold. However, the optimal values may not be attained in the dual problems, as proved in [3, Proposition 4.1]. In [14] and [15] the authors showed that in the at-the-money case ($\mathfrak{K} = 1$), with an additional dispersion assumption on the measures μ and ν the optimal values of the dual problems (2.2) are actually attained. More precisely, the following result summarises [14] (see also Section 4 below for more details about the technical assumption):

Assumption 2.1. The support of $\eta := (\mu-\nu)_+$ is given by an interval $[a,b] \subset \mathbb{R}_+$, and the support of $\gamma := (\nu-\mu)_+$ is given by $\mathbb{R}_+ \setminus [a,b]$.

Theorem 2.2. *The set equalities $\underline{\mathcal{D}}(\mu,\nu) = \underline{\mathcal{P}}(\mu,\nu)$ and $\overline{\mathcal{D}}(\mu,\nu) = \overline{\mathcal{P}}(\mu,\nu)$ hold, and the primal optima in* (2.1) *are attained: there exist martingale measures $\mathbb{Q}_L$ and $\mathbb{Q}_U$ in $\mathcal{M}(\mu,\nu)$ such that $\underline{\mathcal{P}}(\mu,\nu) = \mathbb{E}^{\mathbb{Q}_L}|S_{t+\tau} - \mathfrak{K}S_t|$ and $\overline{\mathcal{P}}(\mu,\nu) = \mathbb{E}^{\mathbb{Q}_U}|S_{t+\tau} - \mathfrak{K}S_t|$. Furthermore, under Assumption 2.1, the infimum and supremum in the dual problems* (2.2) *are attained when $\mathfrak{K} = 1$.*

3. No-Arbitrage discretisation of the primal and dual problems

3.1. **No-arbitrage discretisation of the density.** Let $t > 0$ be some given time horizon, S_t the random variable describing the stock price at time t, and μ the law of S_t. Fix $m > 1$ and suppose that we are given a set $\mathrm{x} = (x_1,\dots,x_m) \in \mathbb{R}_+^m$ of points $0 < x_1 < x_2 < ... < x_m$ in the support of μ, and a discrete distribution q with atom q_i at the point x_i sampled from μ (for example if μ admits density f then one can take $q_i = f(x_i)/\sum_{i=1}^m f(x_i)$). We wish to find a discrete distribution p, close to q, matching the first $l \le m$ moments of μ, in particular satisfying the 'martingale' condition $\langle \mathrm{p},\mathrm{x}\rangle = 1$. Let $\mathrm{T} : \mathbb{R}_+ \to \mathbb{R}_+^l$ be given by $\mathrm{T}(x) := (x, x^2, ..., x^l)$ and define the moment vector $\mathfrak{T} := \int_{\mathbb{R}_+} \mathrm{T}(x)\mu(\mathrm{d}x) \in \mathbb{R}_+^l$. Such a matching condition is not necessarily consistent with a given set of (European) option prices. In order to ensure that the discrete density re-prices these options, we add a second layer: Borwein, Choksi and Maréchal [4] suggested to recover discrete probability distributions from observed market prices of European Call options by minimising the Kullback-Leibler divergence to the uniform distribution (they also comment that any prior distribution can be chosen).

In particular given the law μ of S_t and a set of European Call option prices $\mathfrak{P} \in \mathbb{R}_+^M$, maturing at t with strikes $K_1 < \cdots < K_M$, we can solve the minimisation problem:

$$\min_{\mathrm{p}\in[0,1]^m:\|\mathrm{p}\|=1} \sum_{i=1}^m p_i \log\left(\frac{p_i}{q_i}\right), \qquad \text{subject to } \left(\left(\langle \mathrm{C}_j(\mathrm{x}), \mathrm{p}\rangle\right)_{j=1,\dots,M}, \left(\langle \mathrm{T}_j(\mathrm{x}), \mathrm{p}\rangle\right)_{j=1,\dots,l}\right) = (\mathfrak{P}, \mathfrak{T}). \tag{3.1}$$

for some prior discrete distribution q, where $\mathrm{C}_j(\mathrm{x}) := ((x_1 - K_j)_+, \dots, (x_m - K_j)_+)$, for $j = 1, \dots, M$, denotes the payoff vector of the options and $\mathrm{T}_j(\mathrm{x}) := (x_1^j, \dots, x_m^j)$ is the j-th moment vector $(j = 1, \dots, l)$.

Definition 3.1. The discrete distribution p is consistent with $\mathfrak{P}$ whenever the following hold:

(i) $\sum_{i=1}^m p_i = 1$;
(ii) $\langle \mathrm{C}_j(\mathrm{x}), \mathrm{p}\rangle = \sum_{i=1}^m p_i \mathrm{C}_j(x_i) = \mathfrak{P}_j$ for all $j = 1, \dots, M$;
(iii) $\langle \mathrm{T}_1(\mathrm{x}), \mathrm{p}\rangle = \langle \mathrm{x}, \mathrm{p}\rangle = \mathfrak{T}_1 = 1$;

Note that the last item in the definition above is nothing else than the martingale condition. It must be noted that if the full marginal distribution μ of S_t is known, any finite subset of European options can be chosen above and the price vector can be defined as $\mathfrak{P} := \int_{\mathbb{R}_+} \mathrm{C}(x)\mu(\mathrm{d}x)$, where $\mathrm{C}(x) := ((x - K_1)_+, \dots, (x - K_M)_+) \in \mathbb{R}_+^M$ for any $x \in \mathbb{R}_+$. In particular the solution to this problem can be obtained as a modification of the solution in [23], which itself is based on arguments by Borwein and Lewis [5, Corollary 2.6]:

$$p_i = \frac{q_i \exp\Big(\langle \lambda^*, (\mathrm{C}(x_i), \mathrm{T}(x_i))\rangle\Big)}{\sum_{i=1}^m q_i \exp\Big(\langle \lambda^*, (\mathrm{C}(x_i), \mathrm{T}(x_i))\rangle\Big)}, \quad \text{where} \quad \lambda^* := \operatorname*{argmin}_{\lambda\in\mathbb{R}^{M+l}} \left\{ -\langle \lambda, (\mathfrak{P}, \mathfrak{T})\rangle + \log\left(\sum_{i=1}^m q_i \mathrm{e}^{\langle \lambda, (\mathrm{C}(x_i), \mathrm{T}(x_i))\rangle}\right)\right\}. \tag{3.2}$$

We can now consider the following algorithm:

Algorithm 3.2.

(i) Several choices are possible for the m points $0 < x_1 < ... < x_m$; for instance:
  (a) *Binomial*: Let Σ denote the at-the-money lognormal volatility (for European options maturing at t). Set $\delta := t/(m-1)$, $u := 1 + \left(\mathrm{e}^{\delta\Sigma^2} - 1\right)^{1/2}$, $d := 1 - \left(\mathrm{e}^{\delta\Sigma^2} - 1\right)^{1/2}$, and $x_i := u^{i-1}d^{m-i}$;
  (b) *Gauss-Hermite*: $x_i := \mathrm{e}^{\tilde{x}_i}$, where $\tilde{x}_1, ..., \tilde{x}_m$ are the nodes of an N-point Gauss-Hermite quadrature.
(ii) For the discrete distribution q, we can follow several routes:
  (a) if μ admits a density f_μ, then, for $i = 1, \dots, m$, set $q_i := f_\mu(x_i)/\sum_{j=1}^m f_\mu(x_j)$;
  (b) alternatively, for $i = 1, ..., m$, let $q_i := \mu([x_{i-1}, x_i))$ (with $x_0 = 0$);
(iii) Compute the discretised measure p through (3.2).

Remark 3.3. As pointed out by Tanaka and Toda [23] the choice of discretisation points $x_1, \dots, x_m$ in Algorithm 3.2 is dictated by the discretisation of the integrals $\int_{\mathbb{R}_+} (\mathrm{C}(x), \mathrm{T}(x))\mathrm{d}f_\mu(x) \approx \sum_{i=1}^m w(x_i)(\mathrm{C}(x_i), \mathrm{T}(x_i))f_\mu(x_i)$. The weights $w(\cdot)$ are chosen in accordance with a given quadrature rule; in the case of Algorithm 3.2 the weights are chosen constant $w(x_i) = \left(\sum_{i=1}^m f_\mu(x_i)\right)^{-1}$ for all $i = 1, \dots, m$.

In the particular case where the discretisation nodes and the given strikes satisfy some specific ordering, it is possible to choose a more explicit weighting scheme p consistent with $\mathfrak{P}$ with minimal assumptions on the choice of discretisation:

Lemma 3.4. *Let $m = M + 2$. Consistency with $\mathfrak{P}$ is ensured if both sets of conditions hold:*

(i) $x_1 < K_1$, $x_{M+2} \geq \left(\mathfrak{P}_{M-1}K_M - \mathfrak{P}_M K_{M-1}\right) / \left(\mathfrak{P}_{M-1} - \mathfrak{P}_M\right)$ and $x_{i+1} = K_i$ for $i = 1, \dots, M$;

(ii) $p_{M+2} = \mathfrak{P}_M / (x_{M+2} - K_M)$, $p_1 = 1 - \sum_{j=2}^{M+2} p_j$, *and, for any* $i = M, \dots, 1$,

$$p_{i+1} = \frac{1}{x_{i+1} - K_{i-1}} \left(\mathfrak{P}_{i-1} - \sum_{j=i+2}^{M+2} p_j (x_j - K_{i-1}) \right),$$

with the convention $K_0 := x_1$ *and* $\mathfrak{P}_0 := 1 - x_1$.

Proof. Since the strikes are ordered and the vector $\overline{\mathfrak{P}} := (1, 1, \mathfrak{P}_1, \dots, \mathfrak{P}_M)' \in \mathbb{R}^{M+2}$ satisfies no-arbitrage conditions (in the sense of [8, Theorem 3.1]), the assumption $x_{M+2} \geq (\mathfrak{P}_{M-1} K_M - \mathfrak{P}_M K_{M-1}) / (\mathfrak{P}_{M-1} - \mathfrak{P}_M)$ implies that $x_{M+2} > K_M$. The consistency of the weighting scheme with $\mathfrak{P}$ can be written as $\mathrm{Ap}^{\mathrm{x}} = \overline{\mathfrak{P}}$, with

$$\mathrm{A} := \begin{pmatrix} 1 & 1 & \dots & \dots & 1 \\ x_1 & x_2 & \dots & \dots & x_{M+2} \\ 0 & (x_2 - K_1) & \ddots & \ddots & (x_{M+2} - K_1) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & (x_M - K_{M-1}) & (x_{M+2} - K_{M-1}) \\ 0 & \dots & 0 & 0 & (x_{M+2} - K_M) \end{pmatrix} \quad \text{and} \quad \mathrm{p}^{\mathrm{x}} := \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \\ p_{M+2} \end{pmatrix}.$$

Here A is a real upper triangular matrix, so that the system has a unique solution given in the lemma. It remains to check that the feasible solutions of the system satisfy the additional constraints $p_i \geq 0$ for $i = 1, \dots, M+2$.

Since $x_{M+2} \geq \frac{\mathfrak{P}_{M-1} K_M - \mathfrak{P}_M K_{M-1}}{\mathfrak{P}_{M-1} - \mathfrak{P}_M}$ it follows that

$$p_{M+2} = \frac{\mathfrak{P}_M}{x_{M+2} - K_M} \leq \frac{\mathfrak{P}_{M-1} - \mathfrak{P}_M}{K_M - K_{M-1}} \leq \frac{-K_{M-1} + K_M}{K_M - K_{M-1}},$$

where the last inequality follows from the Put-Call Parity and absence of arbitrage in $\overline{\mathfrak{P}}$. By definition we have

$$p_{M+1} = \frac{1}{K_M - K_{M-1}} (\mathfrak{P}_{M-1} - p_{M+2}(x_{M+2} - K_{M-1})) = \frac{1}{K_M - K_{M-1}} \left(\mathfrak{P}_{M-1} - \mathfrak{P}_M - \mathfrak{P}_M \frac{K_M - K_{M-1}}{x_{M+2} - K_M} \right),$$

and by assumption on x_{M+2} we have that $p_{M+1} \geq 0$. Similarly for p_M we have

$$\begin{aligned} p_M &= \frac{1}{K_{M-1} - K_{M-2}} \left[\mathfrak{P}_{M-2} - p_{M+1}(K_M - K_{M-2}) - p_M (x_{M+2} - K_{M-2}) \right] \\ &= \frac{\mathfrak{P}_{M-2} - \mathfrak{P}_{M-1}}{K_{M-1} - K_{M-2}} - \frac{\mathfrak{P}_{M-1} - p_{M+2}(x_{M+2} - K_M)}{K_M - K_{M-1}} \\ &= \frac{\mathfrak{P}_{M-2} - \mathfrak{P}_{M-1}}{K_{M-1} - K_{M-2}} - \frac{\mathfrak{P}_{M-1} - \mathfrak{P}_M}{K_M - K_{M-1}}. \end{aligned}$$

Proceeding recursively we observe that for $i = 2, \dots, M-1$ we have

$$p_i = \frac{\mathfrak{P}_{i-2} - \mathfrak{P}_{i-1}}{K_{i-1} - K_{i-2}} - \frac{\mathfrak{P}_{i-1} - \mathfrak{P}_i}{K_i - K_{i-1}}.$$

Since the prices $\overline{\mathfrak{P}}$ satisfy no-arbitrage conditions, then $p_i \geq 0$ for $i = 1, \dots, M+2$.

Note that, when $x_{M+2} = \frac{\mathfrak{P}_{M-1} K_M - \mathfrak{P}_M K_{M-1}}{\mathfrak{P}_{M-1} - \mathfrak{P}_M}$,

$$p_{M+2} = \frac{\mathfrak{P}_{M-1} - \mathfrak{P}_M}{K_M - K_{M-1}}, \qquad p_{M+1} = 0, \qquad p_i = \frac{\mathfrak{P}_{i-2} - \mathfrak{P}_{i-1}}{K_{i-1} - K_{i-2}} - \frac{\mathfrak{P}_{i-1} - \mathfrak{P}_i}{K_i - K_{i-1}},$$

for $i = 2, \dots, M$ and the discretisation reduces to $M+1$ points, where $x_{M+1} = K_M$ is discarded. □

Remark 3.5. One could in principle generalise Lemma 3.4 to a construction where $x_{i+1} \in (K_{i-1}, K_i]$ for $i = 1, \dots, M$ (and obviously $x_1 < K_1$ and $x_{M+2} > K_M$). Quick computations however reveal that writing a clear set of sufficient conditions for consistency with $\overline{\mathfrak{P}}$ is notationally cumbersome and practically not particularly enlightening.

3.2. **Balayage-consistent discretisation.** Assume now that there are M European Call options $\mathfrak{P}^{\mathrm{x}}$ available with maturity t and N options $\mathfrak{P}^{\mathrm{y}}$ available with maturity $t+\tau$, and that the measures μ and ν are calibrated to those European options: $\int_{\mathbb{R}_+}(x-K_i^{\mathrm{x}})_+\mu(\mathrm{d}x) = \mathfrak{P}_i^{\mathrm{x}}$ for all $i = 1, \dots, M$ and $\int_{\mathbb{R}_+}(y-K_j^{\mathrm{y}})_+\nu(\mathrm{d}y) = \mathfrak{P}_j^{\mathrm{y}}$ for all $j = 1, \dots, N$. We further assume that $\mathfrak{P}^{\mathrm{x}}$ and $\mathfrak{P}^{\mathrm{y}}$ do not contain arbitrage (in the sense of [8, Theorem 4.2]), otherwise there is no equivalent martingale measure and the primal problem (2.1) is infeasible. For two discretisation meshes $\mathrm{x} = (x_1, \dots, x_m) \in \mathbb{R}_+^m$ and $\mathrm{y} = (y_1, \dots, y_n) \in \mathbb{R}_+^n$, Algorithm 3.2, for instance, produces discrete distributions p^{x} and p^{y} supported on x and y. It is not however guaranteed that the convex ordering of the original measures μ and ν is preserved for p^{x} and p^{y}. The following lemma provides sufficient conditions ensuring this.

Lemma 3.6. *Let* p^{x} *and* p^{y} *be two discrete measures supported on* x *and* y*. The following conditions altogether ensure that* $\mathrm{p}^{\mathrm{x}} \preceq \mathrm{p}^{\mathrm{y}}$*:*

(i) p^{x} *and* p^{y} *have the same mean equal,* $y_1 \le x_1$ *and* $y_n \ge x_m$*;*
(ii) Assumption 2.1 holds with $\mu = \mathrm{p}^{\mathrm{x}}$ *and* $\nu = \mathrm{p}^{\mathrm{y}}$*.*

Remark 3.7. In our framework, the discrete measures p^{x} and p^{y} arise as discretisations of the original measures μ and ν. If p^{x} and p^{y} are to be consistent with $\mathfrak{P}^{\mathrm{x}}$ and $\mathfrak{P}^{\mathrm{y}}$, then the discretisation nodes x and y must be finer than the sets of input strikes, which we can write as

- $x_1 < K_1^{\mathrm{x}}$, $y_1 < K_1^{\mathrm{y}}$, $x_m > K_M^{\mathrm{x}}$ and $y_n > K_N^{\mathrm{y}}$;
- for all $i \in \{1, \dots, M-1\}$, there exists $k_i \in \{2, \dots, m-1\}$ such that $K_i^{\mathrm{x}} \le x_{k_i} \le K_{i+1}^{\mathrm{x}}$, and for all $j \in \{1, \dots, N-1\}$, there exists $k_j \in \{2, \dots, n-1\}$ such that $K_j^{\mathrm{y}} \le y_{k_j} \le K_{j+1}^{\mathrm{y}}$.

Proof of Lemma 3.6. To fix the notations, for any set $A \subset \mathbb{R}$, define $\mathrm{p}^{\mathrm{x}}(A) := \sum_{\{i: x_i \in A\}} p_i^{\mathrm{x}}$ and $\mathrm{p}^{\mathrm{y}}(A) := \sum_{\{j: y_j \in A\}} p_j^{\mathrm{y}}$. In particular for any $z \in [0, y_n]$ we have

$$\mathrm{p}^{\mathrm{y}}([0,z]) = \sum_{\{j \le n: y_j \le z\}} p_j^{\mathrm{y}}, \qquad \text{and} \qquad \mathrm{p}^{\mathrm{x}}([0,z]) = \sum_{\{i \le m: x_i \le z\}} p_i^{\mathrm{x}}.$$

Define furthermore the function $\delta F : \mathbb{R}_+ \to \mathbb{R}$ by $\delta F(z) := \mathrm{p}^{\mathrm{y}}([0,z]) - \mathrm{p}^{\mathrm{x}}([0,z])$. Assumption 2.1 with $\mu = \mathrm{p}^{\mathrm{x}}$ and $\nu = \mathrm{p}^{\mathrm{y}}$, together with the boundary conditions in Lemma 3.6(i), can therefore be written as

$$\begin{cases} \eta(A) = (\mathrm{p}^{\mathrm{x}}(A) - \mathrm{p}^{\mathrm{y}}(A))_+ > 0, & \text{for any } A \subseteq [a,b], \\ \gamma(A) = (\mathrm{p}^{\mathrm{y}}(A) - \mathrm{p}^{\mathrm{x}}(A))_+ > 0, & \text{for any } A \subseteq [0, y_n] \setminus [a,b]. \end{cases} \tag{3.3}$$

Define now the functions $f_{\mathrm{x}}, g_{\mathrm{y}} : \mathbb{R}_+ \to \mathbb{R}_+$ by $f_{\mathrm{x}}(K) := \langle \mathrm{p}^{\mathrm{x}}, (\mathrm{x}-K)_+ \rangle$ and $g_{\mathrm{y}}(K) := \langle \mathrm{p}^{\mathrm{y}}, (\mathrm{y}-K)_+ \rangle$. By [2, Chapter 2, Definition 2.1.6], the balayage condition $\mathrm{p}^{\mathrm{x}} \preceq \mathrm{p}^{\mathrm{y}}$ will be satisfied as soon as f_{x} and g_{y} are convex and $f_{\mathrm{x}}(K) \le g_{\mathrm{y}}(K)$ for all $K \in \mathbb{R}_+$. Since p^{x} and p^{y} have non-negative components, the functions f_{x} and g_{y} are clearly convex, so we are left to prove that $f_{\mathrm{x}}(\cdot) \le g_{\mathrm{y}}(\cdot)$ on $\mathbb{R}_+$.

We first show that the conditions $y_1 \le x_1$ and $x_m \le y_n$ are necessary. If $x_1 < y_1$, then $f_{\mathrm{x}}(K) = g_{\mathrm{y}}(K) = 1-K$ for all $K \in [0, x_1]$. Also, for any $K \in (x_1, x_2 \wedge y_1)$, $g_{\mathrm{y}}(K) = 1-K$ and $f_{\mathrm{x}}(K) = \sum_{i=2}^m p_i^{\mathrm{x}}(x_i - K) =$

$(1-K)+p_1^{\mathrm{x}}(K-x_1) > (1-K)$, which violates the balayage order. Similarly, if $x_m > y_n$, let $i^* := \inf\{i : x_i \geq y_n\}$; then for any $K \in (y_n, x_m)$, $g_{\mathrm{y}}(K) = 0$ and $f_{\mathrm{x}}(K) = \sum_{i=i^*}^m p_i^{\mathrm{x}}(x_i - K)$ which yields the conclusion.

Introduce now the function $G : [0, y_n] \to \mathbb{R}$ as

$$\begin{aligned} G(K) := g_{\mathrm{y}}(K) - f_{\mathrm{x}}(K) &= \sum_{\{j:y_j>K\}} p_j^{\mathrm{y}}(y_j - K) - \sum_{\{i:x_i>K\}} p_i^{\mathrm{x}}(x_i - K) \\ &= \left(\sum_{\{i:x_i>K\}} p_i^{\mathrm{x}} - \sum_{\{j:y_j>K\}} p_j^{\mathrm{y}} \right) K + \sum_{\{j:y_j>K\}} p_j^{\mathrm{y}} y_j - \sum_{\{i:x_i>K\}} p_i^{\mathrm{x}} x_i. \end{aligned} \tag{3.4}$$

The function G is piecewise linear on $[0, y_n]$, not differentiable at the points $\{x_i\}_{1\leq i\leq m}$ and $\{y_j\}_{1\leq j\leq n}$, attains its maximum and minimum on $[0, y_n]$ and $G(0) = G(y_n) = 0$. The lemma then follows if $G(K) \geq 0$ for all $K \in [0, y_n]$. Let $\partial^+ G$ denote its right derivative on $[0, y_n]$ (with the convention $\partial^+ G(y_n) = 0$). The positivity of G together with the boundary conditions at 0 and y_n are therefore equivalent to the following:

$$\partial^+ G(K) \geq 0, \quad \text{for all } K \leq K^* \qquad \text{and} \qquad \partial^+ G(K) \leq 0, \quad \text{for all } K > K^*,$$

where $K^* := \arg\max_K G(K)$. From (3.4), we can therefore write, for any $K \in [0, y_n]$,

$$\partial^+ G(K) = \sum_{\{i:x_i>K\}} p_i^{\mathrm{x}} - \sum_{\{j:y_j>K\}} p_j^{\mathrm{y}} = \sum_{\{j:y_j\leq K\}} p_j^{\mathrm{y}} - \sum_{\{i:x_i\leq K\}} p_i^{\mathrm{x}} = \delta F(K),$$

with δF defined above. Since δF is piecewise constant, it remains to show that it admits a maximum and a minimum attained on unique subsets of $[0, y_n]$. Note first that $\delta F(0) = 0 = \delta F(y_n)$, and that the image of $[0, y_n]$ by δF is exactly $[-1, 1]$. For any $z \in [0, a) \cup (b, y_n]$, $\delta F(z) = \gamma([0, z])$, and therefore δF is an increasing function on $[0, a) \cup (b, y_n]$. On $[a, b]$, however, one has

$$\delta F(z) = \mathrm{p}^{\mathrm{y}}([0, z]) - \mathrm{p}^{\mathrm{x}}([0, z]) = \mathrm{p}^{\mathrm{y}}([a, z]) - \mathrm{p}^{\mathrm{x}}([a, z]) + \mathrm{p}^{\mathrm{y}}([0, a)) - \mathrm{p}^{\mathrm{x}}([0, a)) = -\eta([a, z]) + \gamma([0, a)).$$

Since $\gamma([0, a)) > 0$ and η is a measure on $[a, b]$, then δF is decreasing on $[a, b]$. Therefore there exists a unique quadruplet (c, d, e, f) such that $[c, d] \subset [0, y_n]$, $[e, f] \subset [0, y_n]$ and $[c, d] = \arg\max_z \delta F(z)$ and $[e, f] = \arg\min_z \delta F(z)$. It therefore follows that $\delta F(\cdot) \geq 0$ on $[0, z^*)$ and $\delta F(\cdot) \leq 0$ on $[z^*, y_n]$, for some $z^* \in (d, e)$. □

Remark 3.8. Note that the dispersion Assumption 2.1 intuitively implies that the variance of the underlying price process increases with maturity, which is the case for all stochastic volatility models. Lemma 3.6 provides general conditions to ensure that convex order is preserved under discretisation of continuous measures. These are however difficult to verify analytically for general distributions, or even for those obtained by Algorithm 3.2. In Figure 1 below, we provide numerical evidence that the convex order is preserved for these distributions.

3.3. **Primal and dual formulation.** We focus here on the discretisation of the primal and dual formulations for the upper bound, and note that an analogous formulation holds for the lower bound. We use Algorithm 3.2 to approximate the random variables S_t and $S_{t+\tau}$ by discrete random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$ with distributions p^{x} and p^{y} supported on $\mathrm{x} \in \mathbb{R}_+^m$ and $\mathrm{y} \in \mathbb{R}_+^n$. The linear programme for the primal problem (2.1) then reads

$$\begin{aligned} &\overline{\mathcal{P}}(\mu, \nu) := \max_{\zeta \in \mathscr{M}_+^{m,n}} \sum_{i,j} \zeta_{i,j} |y_j - \mathfrak{K} x_i|, \\ &\text{subject to } (\|\zeta_{i,\cdot}\|)_{i=1,\dots,m} = \mathrm{p}^{\mathrm{x}}, \quad (\|\zeta_{\cdot,j}\|)_{j=1,\dots,n} = \mathrm{p}^{\mathrm{y}}, \quad (\langle \zeta_{i,\cdot}, (x_i - \mathrm{y}) \rangle)_{i=1,\dots,m} = \mathbf{0} \in \mathbb{R}^m, \end{aligned} \tag{3.5}$$

where $\mathscr{M}_+^{m,n}$ denotes the set of matrices of size $m \times n$ with non-negative entries. For the dual problem, denote the Call option price (with strike K and maturity t) on $\widetilde{S}_t$ by $\widetilde{C}(t, K) := \mathbb{E}(\widetilde{S}_t - K)_+ = \sum_{i=i^*}^m (x_i - K) p_i^{\mathrm{x}}$, where

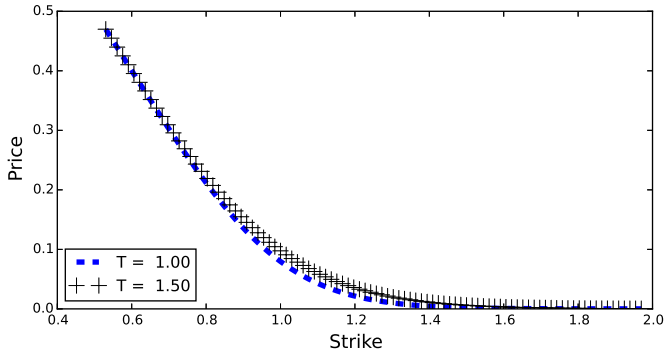


FIGURE 1. Call prices with maturities $t = 1$ (dashed) and $t+\tau = 1.5$ (crosses) computed using the discretised densities of the Black-Scholes model, with parameter $\sigma = 0.2$. As a consistency check, the Call prices with maturity $t + \tau$ are strictly greater than those with maturity t, and both functions are convex.

$i^* := \inf\{1 \leq i \leq m : x_i > K\}$ and $\widetilde{C}(t, K) = 0$ if $x_m \leq K$. The next result is essential for the discretisation of the dual problem, and follows by simple, yet careful, manipulations of telescopic sums.

Lemma 3.9. *Let $l \in \mathbb{N}$ and $\mathrm{K} = (K_1, \ldots, K_l)$. For any one-dimensional real random variable Z, the following representation holds almost surely:*

$$\varphi(Z) = \varphi(K_1) + \mathrm{D}\varphi(K_1)\,(Z - K_1)_+ + \sum_{i=2}^{l-1} \left(\mathrm{D}\varphi(K_i) - \mathrm{D}\varphi(K_{i-1})\right)(Z - K_i)_+ , \tag{3.6}$$

for any continuous function φ, where the forward finite-difference operator D is defined as

$$\mathrm{D}\varphi(K_i) := \frac{\varphi(K_{i+1}) - \varphi(K_i)}{K_{i+1} - K_i}, \qquad \text{for } i = 1, \ldots, l-1.$$

Let now $Z = \widetilde{S}_t$, $\varphi \equiv \psi_0$, and consider the vector of strikes $\mathrm{K}^{\mathrm{x}} \in \mathbb{R}_+^M$. Equality (3.6) can then be rewritten as $\psi_0(\widetilde{S}_t) = w_0^{\mathrm{x}} + w_1^{\mathrm{x}}(\widetilde{S}_t - K_1^{\mathrm{x}})_+ + \sum_{i=2}^{M-1} w_i^{\mathrm{x}}(\widetilde{S}_t - K_i^{\mathrm{x}})_+$, with $l = M$ and where the weights w^{x} read

$$w_0^{\mathrm{x}} := \psi_0(K_1^{\mathrm{x}}), \qquad w_1^{\mathrm{x}} := \mathrm{D}\psi_0(K_1^{\mathrm{x}}), \qquad w_i^{\mathrm{x}} := \mathrm{D}\psi_0(K_i^{\mathrm{x}}) - \mathrm{D}\psi_0(K_{i-1}^{\mathrm{x}}), \text{ for } i = 2, \ldots, M-1,$$

so that

$$\mathbb{E}\psi_0(\widetilde{S}_t) = w_0^{\mathrm{x}} + w_1^{\mathrm{x}}\mathbb{E}\left(\widetilde{S}_t - K_1^{\mathrm{x}}\right)_+ + \sum_{i=2}^{M-1} w_i^{\mathrm{x}}\widetilde{C}(t, K_i^{\mathrm{x}}).$$

Likewise, for $Z = \widetilde{S}_{t+\tau}$, $\varphi \equiv \psi_1$ and $\mathrm{K}^{\mathrm{y}} \in \mathbb{R}_+^N$, an analogous formulation holds at time $t + \tau$:

$$\mathbb{E}\psi_1(\widetilde{S}_{t+\tau}) = w_0^{\mathrm{y}} + w_1^{\mathrm{y}}\mathbb{E}\left(\widetilde{S}_{t+\tau} - K_1^{\mathrm{y}}\right)_+ + \sum_{i=2}^{N-1} w_i^{\mathrm{y}}\widetilde{C}(t+\tau, K_i^{\mathrm{y}}),$$

with with the identification (from (3.6)) $l = N$ and where the weights w^{y} read

$$w_0^{\mathrm{y}} := \psi_0(K_1^{\mathrm{y}}), \qquad w_1^{\mathrm{y}} := \mathrm{D}\psi_0(K_1^{\mathrm{y}}), \qquad w_i^{\mathrm{y}} := \mathrm{D}\psi_0(K_i^{\mathrm{y}}) - \mathrm{D}\psi_0(K_{i-1}^{\mathrm{y}}), \text{ for } i = 2, \ldots, N-1.$$

Define the set $\mathcal{X} := \{(x, y) : x \in \mathrm{Supp}(\widetilde{S}_t), y \in \mathrm{Supp}(\widetilde{S}_{t+\tau})\}$, and assume from now on that K_1^{x} and K_1^{y} are such that $\widetilde{S}_t \geq K_1^{\mathrm{x}}$ and $\widetilde{S}_{t+\tau} \geq K_1^{\mathrm{y}}$ almost surely (equivalently, $K_1^{\mathrm{x}} \leq x_1$ and $K_1^{\mathrm{y}} \leq y_1$), so that the martingale

property (ensured via Algorithm 3.2) yields $\mathbb{E}(\widetilde{S}_t - K_1^{\mathrm{x}})_+ = 1 - K_1^{\mathrm{x}}$ and $\mathbb{E}(\widetilde{S}_{t+\tau} - K_1^{\mathrm{y}})_+ = 1 - K_1^{\mathrm{y}}$. The dual problem (2.2) then reads

(3.7)
$$\overline{\mathcal{D}}(\mu,\nu) = \min\left\{ v + w_1^{\mathrm{x}} + w_1^{\mathrm{y}} + \sum_{i=2}^{M-1} w_i^{\mathrm{x}} \widetilde{C}(t, K_i^{\mathrm{x}}) + \sum_{i=2}^{N-1} w_i^{\mathrm{y}} \widetilde{C}(t+\tau, K_i^{\mathrm{y}}) : (\mathrm{w}^{\mathrm{x}}, \mathrm{w}^{\mathrm{y}}, \delta) \in \mathbb{R}^{M+N} \times C_b(\mathbb{R}_+) \right\},$$

with $\mathrm{w}^{\mathrm{x}} = (w_0^{\mathrm{x}}, \dots, w_{M-1}^{\mathrm{x}})$, $\mathrm{w}^{\mathrm{y}} = (w_0^{\mathrm{y}}, \dots, w_{N-1}^{\mathrm{y}})$, subject to the constraints

$$\begin{cases} v + w_1^{\mathrm{x}} x + w_1^{\mathrm{y}} y + \sum_{i=2}^{M-1} w_i^{\mathrm{x}} (x - K_i^{\mathrm{x}})_+ + \sum_{i=2}^{N-1} w_i^{\mathrm{y}} (y - K_i^{\mathrm{y}})_+ + \delta(x)(y-x) \;\geq\; |y - \mathfrak{K} x|, \quad \text{for all } (x,y) \in \mathcal{X}, \\ w_0^{\mathrm{x}} + w_0^{\mathrm{y}} - w_1^{\mathrm{x}} K_1^{\mathrm{x}} - w_1^{\mathrm{y}} K_1^{\mathrm{y}} \;=\; v. \end{cases}$$

The dual problem here is semi-infinite dimensional since the minimisation is performed over the finite-dimensional vectors w^{x} and w^{y}, but also over the space of continuous and bounded functions δ on $\mathbb{R}_+$ (it is enough to consider only continuous and bounded functions instead of bounded measurable functions as pointed out in [3, Section 1.5]). The importance of incorporating the martingale conditions into the discretisation is critical. This is easily seen in the following example for the primal problem, which also yields an issue for the dual. Suppose that S_t can take value 0.75 or 1.25 each with 50% probability and $S_{t+\tau}$ can take value 0.5 or 1.5 each with 50% probability. Note that $\mathbb{E}(S_t) = \mathbb{E}(S_{t+\tau}) = 1$. We consider the primal problem. The constraints $\|\zeta_{i,\cdot}\| = \mu_i$ and $(\langle \zeta_{i,\cdot}, (x_i - \mathrm{y}) \rangle)_i = \mathbf{0}$ fully determine the probabilities $\zeta_{1,1} = \zeta_{2,2} = 3/8$ and $\zeta_{1,2} = \zeta_{2,1} = 1/8$. The final constraints $\|\zeta_{\cdot,j}\| = \nu_j$ are only true if $\nu_1 = \nu_2 = 1/2$ or $\mathbb{E}(S_{t+\tau}) = 1$. Otherwise, there will be no solution to this LP. This stresses the importance of a consistent no-arbitrage discretisation of the problem.

3.3.1. *Approximation of the dual.* In order to reduce the dual problem (3.7) to a purely finite-dimensional problem, we further add a layer of discretisation for the continuous and bounded delta hedges δ. Similar to [10], fix $M_b \in \mathbb{N}$ and a finite-dimensional basis $(\phi_i)_{i=1,\dots,M_b}$ on $C_b(\mathbb{R}_+)$, and let $\mathrm{w}^b := (w_1^b, \dots, w_{M_b}^b)$ be a vector in $\mathbb{R}^{M_b}$; define then the discretised hedge $\widetilde{\delta} : \mathbb{R}_+ \to \mathbb{R}$ as

$$\widetilde{\delta}(x) := \sum_{i=1}^{M_b} w_i^b \phi_i(x), \tag{3.8}$$

so that the new (discretised) dual problem now has the following finite-dimensional form:

$$\overline{\mathcal{D}}_b(\mu,\nu) = \min\left\{ v + w_1^{\mathrm{x}} + w_1^{\mathrm{y}} + \sum_{i=2}^{M-1} w_i^{\mathrm{x}} \widetilde{C}(t, K_i^{\mathrm{x}}) + \sum_{i=2}^{N-1} w_i^{\mathrm{y}} \widetilde{C}(t+\tau, K_i^{\mathrm{y}}) : (\mathrm{w}^{\mathrm{x}}, \mathrm{w}^{\mathrm{y}}, \mathrm{w}^b) \in \mathbb{R}^{M+N+M_b} \right\}, \tag{3.9}$$

subject to the constraints

$$\begin{cases} v + w_1^{\mathrm{x}} x + w_1^{\mathrm{y}} y + \sum_{i=2}^{M-1} w_i^{\mathrm{x}} (x - K_i^{\mathrm{x}})_+ + \sum_{i=2}^{N-1} w_i^{\mathrm{y}} (y - K_i^{\mathrm{y}})_+ + \widetilde{\delta}(x)(y-x) \;\geq\; |y - \mathfrak{K} x|, \quad \text{for all } (x,y) \in \mathcal{X}, \\ w_0^{\mathrm{x}} + w_0^{\mathrm{y}} - w_1^{\mathrm{x}} K_1^{\mathrm{x}} - w_1^{\mathrm{y}} K_1^{\mathrm{y}} \;=\; v. \end{cases}$$

4. Primal solution for the at-the-money case

In [14], Hobson and Klimmek derived the lower bound optimal martingale transport plan for the at-the-money ($\mathfrak{K} = 1$) forward-start straddle. Let $\Delta(z) := \int_0^z f_\nu(u)\mathrm{d}u - \int_0^z f_\mu(u)\mathrm{d}u$ for all $z \geq 0$; then Assumption 2.1—crucial in Hobson and Klimmek's analysis—is equivalent to Δ having a single maximiser [7, Lemma 5.1]. This assumption imposes constraints on the tail behaviour of the difference between the two laws μ and ν, and is clearly satisfied, for instance, in the Black-Scholes case.

4.1. **Structure of the transport plan.** The key risk for an at-the-money forward-start straddle is that a long position is equivalent to being short the kurtosis of the conditional distribution of the underlying asset (see for example in the introduction of [15]). Therefore to produce the lowest possible price it seems reasonable to require a transport plan that maximises the kurtosis of the conditional distribution. This is indeed the structure of the solution in [14]. We leave as much common mass $(\mu \wedge \nu)$ in place and then map the residual mass η on $[a,b]$ to the tails of the distribution γ via two decreasing functions $p : [a,b] \to [0,a]$ and $q : [a,b] \to [b,\infty)$. Using the martingale condition, Hobson and Klimmek [14] derive a system of coupled differential equations for (p,q):

$$p'(x) = \frac{q(x)-x}{q(x)-p(x)} \frac{f_\mu(x)-f_\nu(x)}{f_\mu(p(x))-f_\nu(p(x))}, \qquad q'(x) = \frac{x-p(x)}{q(x)-p(x)} \frac{f_\mu(x)-f_\nu(x)}{f_\mu(q(x))-f_\nu(q(x))}, \tag{4.1}$$

with boundary conditions

$$\begin{aligned} p(b) &= \inf\{x \geq 0 : \gamma([0,x]) > 0\}, \\ q(b) &= \inf\{x \geq 0 : \gamma([0,x]) > \gamma([0,b])\}, \\ p(a) &= \sup\{x \geq 0 : \eta([0,x]) < \eta([0,a])\}, \\ q(a) &= \sup\{x \geq 0 : \gamma([0,x]) < 1\}. \end{aligned}$$

By taking limits, we obtain that $p(a) = a$, $p(b) = 0$ $q(a) = +\infty$ and $q(b) = b$, see also [7, Proposition 5.6].

4.2. **Implementation.** The right-hand side of the equations in (4.1) are undefined at the boundary points. An application of L'Hôpital's rule shows that $\lim_{x\uparrow b} q'(x) = -1$ if $f'_\mu(b) \neq f'_\nu(b)$, which is a reasonable assumption in practice, as will be illustrated in Section 5. On the other hand $\lim_{x\uparrow b} p'(x)$ depends on the marginal measures μ and ν. For instance, in the lognormal example in Section 5, we find that $p'(x) = \mathcal{O}\left(\mathrm{e}^{\alpha(\log p(x))^2}\right)$ for some $\alpha > 0$ as x tends to b from below, and $\lim_{x\uparrow b} p'(x) = -\infty$ (see for example Figure 2(b)). On the other hand if for example $f_\mu(0) \neq f_\nu(0)$ or $f'_\mu(0) \neq f'_\nu(0)$ then $\lim_{x\uparrow b} p'(x) = 0$.

In order to circumvent these issues so that we can apply the Runge-Kutta method to solve these ODEs, we introduce the following pre-processing step: fix a small $\varepsilon > 0$ (in our implementations, we choose $\varepsilon = 0.001$), integrate both sides of (4.1) over $[b-\varepsilon, b]$ and then approximate the right-hand side by using the rectangle rule for the integral with the unknown values $p^* := p(b-\varepsilon)$ and $q^* := q(b-\varepsilon)$. This yields the following simultaneous equations for p^* and q^* which we solve numerically:

$$\begin{aligned} p^* &= -\varepsilon \frac{q^*-b+\varepsilon}{q^*-p^*} \frac{f_\mu(b-\varepsilon)-f_\nu(b-\varepsilon)}{f_\mu(p^*)-f_\nu(p^*)}, \\ q^* &= b - \varepsilon \frac{b-\varepsilon-p^*}{q^*-p^*} \frac{f_\mu(b-\varepsilon)-f_\nu(b-\varepsilon)}{f_\mu(q^*)-f_\nu(q^*)}. \end{aligned}$$

These equations can easily be reduced into one root search; the first equation gives

$$q^* = \frac{(p^*)^2 \left[f_\mu(p^*)-f_\nu(p^*)\right] + \varepsilon(b-\varepsilon)\left[f_\mu(b-\varepsilon)-f_\nu(b-\varepsilon)\right]}{p^* \left[f_\mu(p^*)-f_\nu(p^*)\right] + \varepsilon \left[f_\mu(b-\varepsilon)-f_\nu(b-\varepsilon)\right]},$$

which in turn can be plugged into the second equation to solve for p^*. The pair (p,q) in (4.1) is then solved for $x \in [a, b-\varepsilon]$ using standard Runge-Kutta methods with the new boundary conditions $p(b-\varepsilon) = p^*, q(b-\varepsilon) = q^*$. The lower bound for the at-the-money forward-start straddle price is then given using the optimal conditional

density $\rho^*(Y=y|X=x)$ given in [14, page 199]:

$$\rho^*(Y=y|X=x) = \begin{cases} \dfrac{f_\eta(x)}{f_\mu(x)} \dfrac{q(x)-x}{q(x)-p(x)} \mathbf{1}_{\{y=p(x)\}}, & \text{if } y<x, \\ \left(1 - \dfrac{f_\eta(x)}{f_\mu(x)}\right), & \text{if } y=x, \\ \dfrac{f_\eta(x)}{f_\mu(x)} \dfrac{x-p(x)}{q(x)-p(x)} \mathbf{1}_{\{y=q(x)\}}, & \text{if } y>x, \end{cases}$$

and hence straightforward computations yield

$$\begin{aligned} \mathbb{E}(|Y-X|) &= \int_{\mathbb{R}} \mathbb{E}(|Y-X||X=x) f_\mu(x) \mathrm{d}x = \int_{\mathbb{R}} \int_{\mathbb{R}} \rho^*(Y=y|X=x) f_\mu(x) |y-x| \mathrm{d}y \mathrm{d}x \\ &= \int_{\mathbb{R}} \int_{-\infty}^{x} f_\eta(x) \frac{q(x)-x}{q(x)-p(x)} \mathbf{1}_{\{y=p(x)\}} |y-x| \mathrm{d}y \mathrm{d}x + \int_{\mathbb{R}} \int_{x}^{+\infty} f_\eta(x) \frac{x-p(x)}{q(x)-p(x)} \mathbf{1}_{\{y=q(x)\}} |y-x| \mathrm{d}y \mathrm{d}x \\ &= \int_a^b \frac{2(x-p(x))(q(x)-x)}{q(x)-p(x)} f_\eta(x) \mathrm{d}x. \end{aligned} \tag{4.2}$$

5. Numerical analysis of the no-arbitrage bounds

We now illustrate the numerical methods developed in Sections 3 and 4 on the Black-Scholes and the Heston models. These examples involve forward-start option prices (and forward implied volatility smile), which we now quickly recall. In the Black-Scholes model, the dynamics of the stock price process under the risk-neutral measure are given by $\mathrm{d}S_t = S_t \Sigma \mathrm{d}W_t$, $S_0 = 1$, where $\Sigma > 0$ represents the instantaneous volatility and W is a standard Brownian motion. The no-arbitrage price of the Call option at time zero then reads $\mathrm{BS}(\tau, K, \Sigma) := \mathbb{E}\left(S_\tau - K\right)^+ = \mathcal{N}\left(d_+\right) - \mathrm{e}^k \mathcal{N}\left(d_-\right)$, with $d_\pm := -\frac{k}{\Sigma\sqrt{\tau}} \pm \frac{1}{2}\Sigma\sqrt{\tau}$, where $\mathcal{N}$ is the standard normal distribution function. Since the increments of the stock price process are stationary and independent, the forward-start option with payoff $(S_{t+\tau} - KS_t)^+$ with $t, \tau > 0$ is worth $\mathrm{BS}(\tau, K, \Sigma)$. For a given market or model price $C^{\mathrm{obs}}(t, \tau, K)$ of the option at strike K, forward-start date t and maturity τ, the forward implied volatility smile $\sigma_{t,\tau}(K)$ is then defined as the unique solution to $C^{\mathrm{obs}}(t, \tau, K) = \mathrm{BS}(\tau, k, \sigma_{t,\tau}(K))$.

In the following two subsections, we shall consider the observed vanilla Call option prices as computed from the Black-Scholes and the Heston model. We shall discretise the corresponding dual problems following Section 3.3, by considering the compact interval $[0, 5]$ for the supports of the discretised random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$ with $m = n = 500$ discretisation points; the vectors of observed strikes are taken as $\mathrm{K^x} = \mathrm{K^y} = \{0.3, 0.4, 0.5, \ldots, 2\}$.

5.1. Application to the Black-Scholes model.

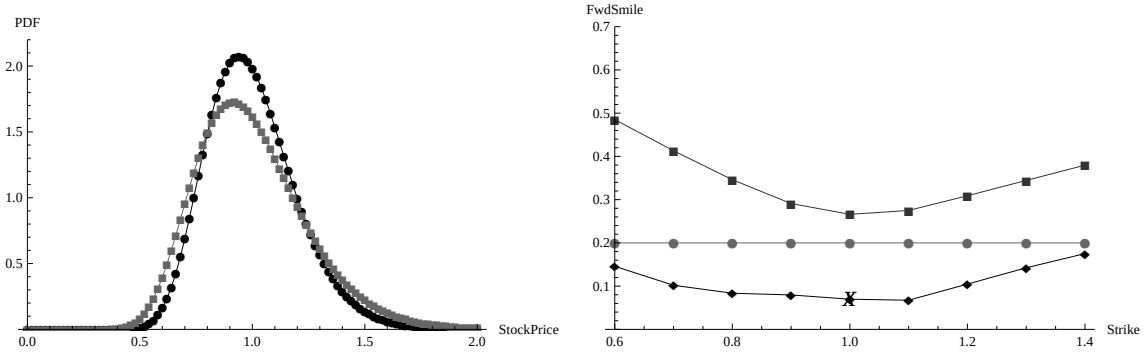
Let $\mathcal{N}(\overline{m}, \Sigma^2)$ denote the Gaussian distribution with mean $\overline{m}$ and variance Σ^2, and assume that the random variables S_t and $S_{t+\tau}$ are distributed according to

$$\log(S_t) \sim \mathcal{N}\left(-\frac{1}{2}\Sigma^2 t, \Sigma^2 t\right) \qquad \text{and} \qquad \log(S_{t+\tau}) \sim \mathcal{N}\left(-\frac{1}{2}\Sigma^2 (t+\tau), \Sigma^2 (t+\tau)\right).$$

Clearly a candidate martingale coupling is the Black-Scholes model with volatility Σ and starting at $S_0 = 1$; in this case the forward volatility, i.e. the implied volatility computed from the forward-start option, is constant and also equal to Σ. In Figure 2(a), we consider the values $\Sigma = 0.2$, $t = 1$ and $\tau = 0.5$, and plot the distributions of S_t and $S_{t+\tau}$.

In Figure 2(b) we plot the lower and upper bounds for the forward implied volatility smile computed from the discretisation of the dual problems (2.2) via (3.9). The lower bound at-the-money case using the Hobson-Klimmek solution and the LP dual solution are virtually identical (6.95% vs 6.98%), illustrating the consistency of the two approaches. Note that even in this simple case the range of possible forward smiles consistent with

the two marginal laws is wide. The magnitude of the bounds produced is similar to bounds obtained in [10] for cliquet options which are very closely related to forward-start straddles. This behaviour can be explained intuitively as no conditional instruments are used to hedge the straddle, hence there is no restriction on the conditional probabilities between times t and $t+\tau$. The only restriction is that the two marginal laws at those times are placed in convex order, producing a very large class of feasible martingale measures.



(a) Densities.

(b) Robust bounds via the dual problem.

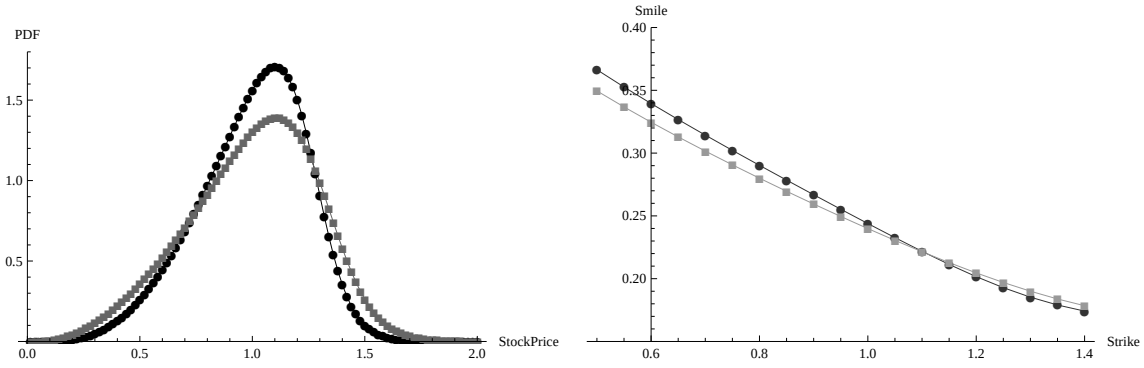
FIGURE 2. (a) the circles represent the one-year lognormal density and the squares the 1.5-year lognormal density. (b) the circles represent the constant Black-Scholes forward volatility Σ consistent with the marginals. The squares and the diamonds are the lower and upper bounds found by solving the dual problems (2.2) via (3.9); the X cross is the Hobson-Klimmek solution (4.2) for the lower bound at-the-money case.

5.2. **Application to the Heston model.** The marginal distributions for expiries $t = 1$ and $t + \tau = 1.5$ are now generated according to the Heston stochastic volatility model [11], in which the stock price process is the unique strong solution to the stochastic differential equation

$$\begin{aligned} \mathrm{d}S_t &= S_t\sqrt{V_t}\mathrm{d}W_t, & S_0 = 1, \\ \mathrm{d}V_t &= \kappa\left(\theta - V_t\right)\mathrm{d}t + \xi\sqrt{V_t}\mathrm{d}Z_t, & V_0 = v > 0, \end{aligned} \tag{5.1}$$

where W and Z are two one-dimensional standard Brownian motions with $\mathrm{d}\langle W, Z\rangle_t = \rho\mathrm{d}t$, and $\kappa, \theta, \xi > 0$ and $\rho \in [-1, 1]$. We consider here the following values: $v = \theta = 0.07$, $\kappa = 1$, $\xi = 0.4$ and $\rho = -0.8$. The (spot) implied volatility smiles and corresponding densities are displayed in Figure 3. Figure 4 shows the Heston forward smile consistent with the marginals (computed using the inverse Fourier transform representation and a simple root search method) and the lower and upper bounds for the forward smile, from the discretised version (3.9) of the dual problems (2.2). For the discretisation of the delta hedge (3.8), we consider the monomials $(\phi_1, \phi_2, \phi_3)(x) \equiv (1, x, x^2)$. As in the Black-Scholes case, the Hobson-Klimmek solution (7.77%) and the dual solution (7.80%) for the lower-bound at-the-money volatility are virtually identical. Figure 5(a) shows the payoff of the option prices in the super-hedge: one enters into positions with long convexity for the 1.5-year maturity and short convexity for the 1-year maturity.

In both examples the range of forward smiles consistent with the marginal laws is large. Using European options to 'lock-in' (replicate) forward volatility or hedge forward volatility dependent claims seems illusory. Forward-start options should be seen as fundamental building blocks for exotic pricing and not decomposable



(a) Densities. (b) Spot implied volatility.

FIGURE 3. (a) Circles (squares) represents the 1 year (1.5 year) marginal densities. (b) Circles (squares) represent the corresponding spot implied volatilities.

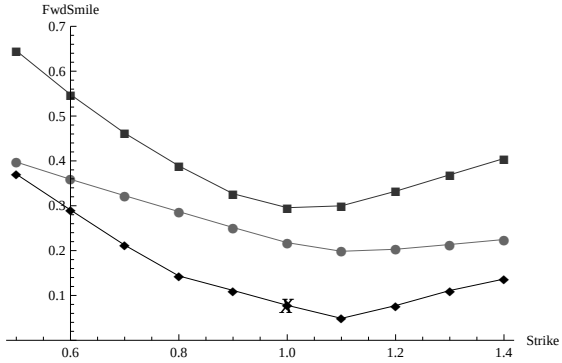
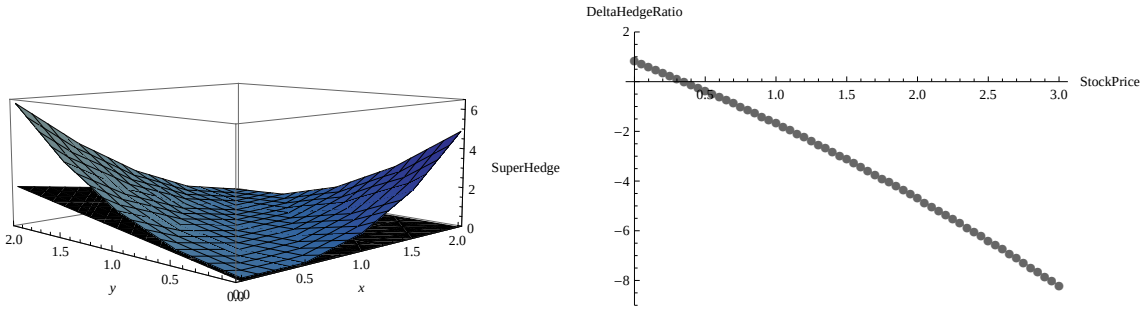


FIGURE 4. The circles represent the Heston forward volatility consistent with the marginals, and the squares and diamonds stand far the lower and upper bounds found by solving the LP problem in Section 3, and X is the primal Hobson-Klimmek solution for the lower bound at-the-money case (Section 4).



(a) Super-hedging portfolio and forward-start payoffs (b) Discretised Delta hedge

FIGURE 5. The strike is taken at the money: $\mathfrak{K} = 1$. (a) Super-hedging portfolio (top plot) and forward-start payoffs (below); (b) Optimal discretised delta hedge $\widetilde{\delta}$ defined in (3.8).

(or approximately decomposable) into European options. Models used for forward volatility dependent exotics should have the capability of calibration to forward-start option prices and at a minimum should produce realistic forward smiles that are consistent with trader expectations and observable prices.

5.3. **A note on the discretisation methodology.** In both examples above, the discretised random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$ were supported on 500 points. This choice was arbitrary, and it is natural to question it. In the semi-infinite case, the absence of duality gap between the primal problem and its dual guarantees [19, Theorem 3.1] the existence of a discretisation, the value of which converges to that of the primal problem as the number of points increases. Rates of convergence have also been obtained for discretisation schemes in semi-infinite programming [19, 21]. Our setting here (Primal problem (2.1) and its dual (2.2)) is however of infinite-dimensional nature, and, to the best of our knowledge, no corresponding result exists (yet!). A general, theoretical, proof is outside the scope of our approach, and we now provide some numerical evidence about the convergence and stability of our discretisation scheme (3.9). We choose two different discretisation grids for the supports of the random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$, following Algorithm 3.2: a uniform grid and a grid consisting of roots of Legendre polynomials, both with the same number of points. Below, Tables 1 and 2 represent the optimal values of the sub-hedging dual problem as the number of discretisation points increases, for different values of the forward-start straddle strike $\mathfrak{K}$. Clearly, refining the discretisation grid produces only minor changes in the optimal values for the at-the-money case ($\mathfrak{K} = 1$) and virtually none for the other cases. Tables 3 and 4 represent the optimal values of the super-hedging dual problem for different numbers of discretisation points and different values of the forward-start straddle strike $\mathfrak{K}$. Refining the discretisation grid yields only minor changes to the optimal values. In contrast to the sub-hedging case, the change in optimal values for the super-hedging problem can be observed for all strikes. Finally, both sub- and super-hedging dual problems seem to be very stable with respect to the discretistation schemes and the number of points.

We would also like to comment on the rate of convergence of the optimal portfolios with respect to the partition refinements. Hobson and Neurberger [15, Section 6.2] obtained explicit expressions for the dual variables (ψ_0, ψ_1, δ), defined in (2.2), in the lognormal case for the at-the-money forward-start straddle ($\mathfrak{K} = 1$):

$$\begin{aligned} \psi_0(x) &= -\xi x \ln x + \xi x \ln\left(A \sinh(\xi^{-1})\right) + x \coth(\xi^{-1}), \\ \psi_1(y) &= \xi \left(y \ln y - y \ln(A/\xi) - y\right), \\ \delta(x) &= -\xi \ln\left(x/A \sinh(\xi^{-1})\right), \end{aligned} \tag{5.2}$$

where the constants A and ξ are such that the expected value of the hedging portfolio $\psi_0(x)+\psi_1(y)+\delta(x)(y-x)$ is minimised under super-hedging constraints. They also showed that the bound achieved by the solution is tight [15, Section 10]. We compute the sup-norm error $\varepsilon(n)$ (n is the number of discretisation points) between solutions of the discretised Dual (3.9) and the Hobson-Neuberger solution (5.2). We visually check the rate of convergence (i.e. the highest exponent r such that $\varepsilon(n) \sim \mathcal{O}(d_n^r)$) with respect to the discretisation mesh size d_n by plotting $\log(\varepsilon(n))/\log(d_n)$ against $\log(d_n)$. The resulting plot is presented in Figure 6 below. As mentioned above, no theoretical rates of convergence exist for infinite-dimensional linear programming problems; in the semi-infinite case, the corresponding plot would be roughly constant.

Strike \ Number of points	75	250	500	1000	2000
0.6	0.4	0.4	0.4	0.4	0.4
0.7	0.3	0.3	0.3	0.3	0.3
0.8	0.2	0.2	0.2	0.2	0.2
0.9	0.1001	0.1	0.1	0.1	0.1
1.0	0.0390	0.0384	0.0384	0.0384	0.0383
1.1	0.1004	0.1004	0.1004	0.1004	0.1004
1.2	0.2	0.2	0.2	0.2	0.2
1.3	0.3	0.3	0.3	0.3	0.3
1.4	0.4	0.4	0.4	0.4	0.4

TABLE 1. Optimal values of the sub-hedging dual problem (using roots of Legendre polynomials).

Strike \ Number of points	75	250	500	1000	2000
0.6	0.4	0.4	0.4	0.4	0.4
0.7	0.3	0.3	0.3	0.3	0.3
0.8	0.2	0.2	0.2	0.2	0.2
0.9	0.10008	0.1	0.1	0.1	0.1
1.0	0.0388	0.0384	0.0384	0.0384	0.0383
1.1	0.1006	0.1004	0.1004	0.1004	0.1004
1.2	0.2	0.2	0.2	0.2	0.2
1.3	0.3	0.3	0.3	0.3	0.3
1.4	0.4	0.4	0.4	0.4	0.4

TABLE 2. Optimal values of the sub-hedging dual problem (using a uniform grid).

Strike \ Number of points	75	250	500	1000	2000
0.6	0.4147	0.4156	0.4157	0.4157	0.4157
0.7	0.3241	0.3255	0.3257	0.3257	0.3257
0.8	0.2394	0.2411	0.2413	0.2413	0.2414
0.9	0.1707	0.1743	0.1745	0.1746	0.1746
1.0	0.1453	0.1485	0.1489	0.1490	0.1490
1.1	0.1786	0.1815	0.1816	0.1817	0.1817
1.2	0.2513	0.2536	0.2538	0.2539	0.2539
1.3	0.3373	0.3394	0.3396	0.3397	0.3397
1.4	0.4292	0.4314	0.4316	0.4316	0.4316

TABLE 3. Optimal values of the super-hedging dual problem (using roots of Legendre polynomials).

Strike \ Number of points	75	250	500	1000	2000
0.6	0.4150	0.4157	0.4157	0.4157	0.4157
0.7	0.3245	0.3256	0.3257	0.3257	0.3257
0.8	0.2397	0.2411	0.2413	0.2413	0.2414
0.9	0.1716	0.1743	0.1746	0.1746	0.1746
1.0	0.1463	0.1487	0.1489	0.1490	0.1490
1.1	0.1795	0.1814	0.1817	0.1817	0.1817
1.2	0.2514	0.2538	0.2539	0.2539	0.2539
1.3	0.3376	0.3395	0.3396	0.3397	0.3397
1.4	0.4300	0.4315	0.4316	0.4316	0.4316

TABLE 4. Optimal values of the super-hedging dual problem (using a uniform grid).

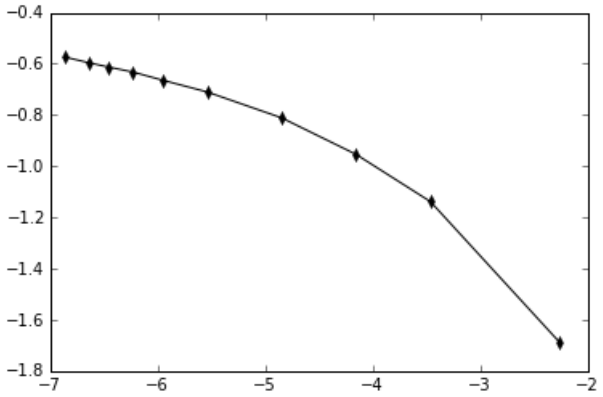


FIGURE 6. Plot of $\log(\varepsilon(n))/\log(d_n)$ against $\log(d_n)$, where $\varepsilon(n)$ is the sup-norm error.

6. Numerical analysis of the transport plans

As mentioned in Section 4, the key risk for the at-the-money forward-start straddle is that a long position is equivalent to being short the kurtosis of the conditional distribution. The solution in the lower bound case (under Assumption 2.1) was detailed in Section 4, where – intuitively – the transport plan maximises the kurtosis of the conditional distribution. In the upper bound case (see [15]) the support of the transport plan is concentrated on a binomial map with no mass being left in place, i.e. all the mass of μ gets mapped to ν via two increasing, continuous and differentiable functions $f, g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $f(x) \leq x \leq g(x)$. Functions f and g must satisfy the system of integral equations [15, Equations (5.19) and (5.20)]

$$\begin{cases} 0 = \displaystyle\int_{g^{-1}(y)}^{f^{-1}(y)} \frac{\big(g(z) - f^{-1}(y)\big)}{g(z) - f(z)} \mathrm{d}z, \\ 1 = \displaystyle\int_{g^{-1}(y)}^{f^{-1}(y)} \frac{1}{g(z) - f(z)} \mathrm{d}z, \end{cases}$$

and if it is possible to find a solution to this system then [15, Lemma 7.1] provides an optimality result in the at-the-money case. Intuitively in this case the solution minimises the kurtosis of the conditional distribution. For

out-of-the-money options the situation is more subtle. As the strike moves further away from the money, a long option position becomes longer the kurtosis of the conditional distribution. Intuitively one would then expect the transport plan to be some combination of the lower and upper at-the-money transport plans discussed above. Using the lognormal example of Section 5, we now numerically solve for the transport plans using the discretised primal problem (3.5) and make qualitative conjectures concerning the structure of the transport plans. The supports of the discretised random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$ (introduced in Section 3.3) are taken respectively as $[0, 10]$ (with $m = 1000$ points) and $[0, 30]$ (with $n = 3000$ points); the vectors of observed strikes are again $\mathrm{K}^{\mathrm{x}} = \mathrm{K}^{y} = \{0.3, 0.4, 0.5, \ldots, 2\}$, and the discrete probabilities in each bucket were obtained by integrating the lognormal density over each bucket.

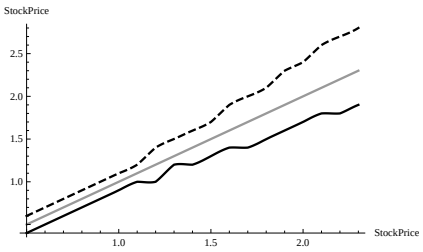
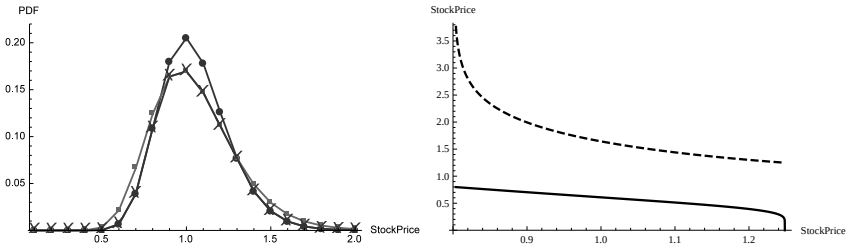


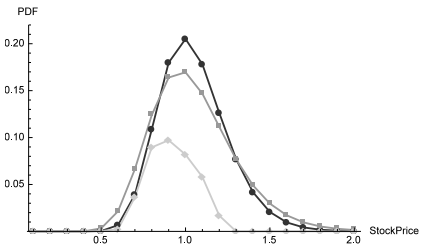
FIGURE 7. The dashed and dark lines are the transport maps for the upper bound at-the-money case and the grey line is the identity. The horizontal and vertical axes are S_t and $S_{t+\tau}$.



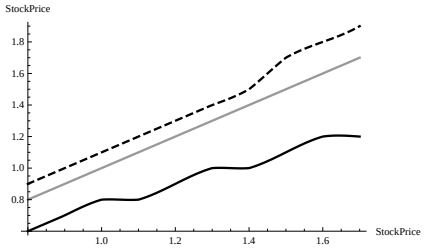
(a) Lower Bound. (b) Transport maps

FIGURE 8. Here the strike is taken at the money $\mathfrak{K} = 1$. (a) Discretisation of the measures μ (circles) and ν (squares) and the amount of mass that must be left in place (X's) in the transport plan for the at-the-money ($\mathfrak{K} = 1$) lower bound case. (b) Transport maps for the residual mass, computed from (4.1) or, equivalently, by solving the primal problem (3.5).

In Figure 7 we compute the transport maps f and g for the at-the-money upper bound case, i.e. the supremum case in (2.1). In this case no mass is left in place in the transport plan; in Figure 8 we plot the transport plan for the at-the-money lower bound case. This lower bound case is in striking agreement with Hobson-Klimmek: as much mass as possible is left in place and the residual mass is mapped to the tails of the distribution via two decreasing functions. Note the agreement with the transport maps in Figure 2(b). In this case the forward volatility is 6.92% matching the Hobson-Klimmek analytical solution and the numerical solution of the dual. Figures 9 and 10 illustrate the transport plan for the upper bound case and strikes $\mathfrak{K} = 0.7$ and $\mathfrak{K} = 0.9$. As the strike decreases from at-the-money, more and more mass is left in place (starting from the left tail), and the residual mass of μ is mapped to ν via two increasing functions; one maps the residual mass to the left tail

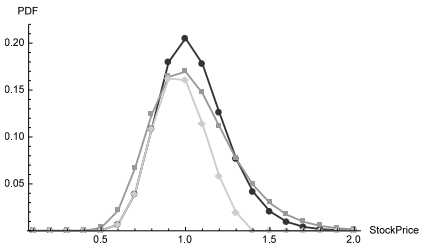


(a) Mass in place $\mathfrak{K} = 0.9$: Upper Bound.

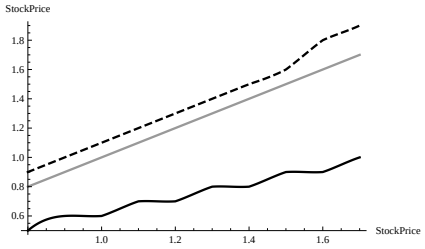


(b) Transport Maps $\mathfrak{K} = 0.9$: Upper Bound.

FIGURE 9. (a) Discretisation of the measures μ (circles), ν (squares) and the amount of mass that must be left in place (diamonds) in the transport plan for the $\mathfrak{K} = 0.9$ upper bound case. (b) Transport maps for the residual mass: the axes are labelled as in Figure 7.

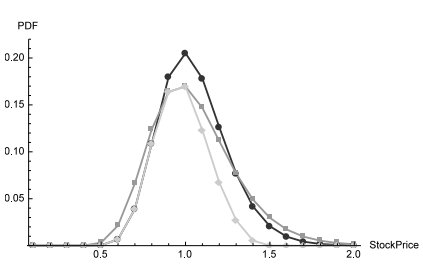


(a) Mass in place $\mathfrak{K} = 0.7$: Upper Bound.

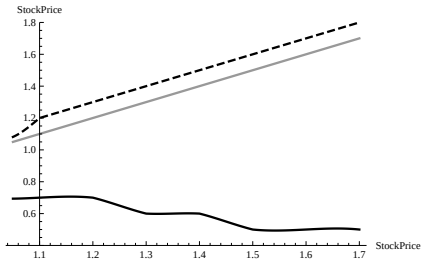


(b) Transport Maps $\mathfrak{K} = 0.7$: Upper Bound.

FIGURE 10. (a) Discretisation of the measures μ (circles), ν (squares) and the amount of mass that must be left in place (diamonds) in the transport plan for the $\mathfrak{K} = 0.7$ upper bound case. (b) Transport maps for the residual mass: the axes are labelled as in Figure 7.



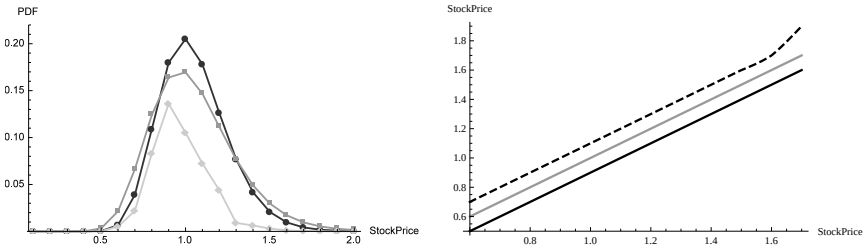
(a) Mass in place $\mathfrak{K} = 1.05$: Lower Bound.



(b) Transport Maps $\mathfrak{K} = 1.05$: Lower Bound.

FIGURE 11. (a) Discretisation of the measures μ (circles), ν (squares) and the amount of mass that must be left in place (diamonds) in the transport plan for the $\mathfrak{K} = 1.05$ lower bound case. (b) Transport maps for the residual mass: the axes are labelled as in Figure 7.

of ν while the other maps the residual mass to the right tail of ν. For strikes greater than at-the-money a mirror-image transport plan emerges where more and more mass is left in place (starting from the right tail) and again the residual mass of μ is mapped to ν via two increasing functions (for brevity we omit the plots).



(a) Mass in place $\mathfrak{K} = 1.3$: lower bound. (b) Transport maps $\mathfrak{K} = 1.3$: lower bound.

FIGURE 12. (a) Discretisation of the measures μ (circles), ν (squares) and the amount of mass left in place (diamonds) in the transport plan for the $\mathfrak{K} = 1.3$ lower bound case. (b) Transport maps for the residual mass: the axes are labelled as in Figure 7.

Figures 11 and 12 illustrate the transport plan for the lower bound case and strikes $\mathfrak{K} = 1.05$ and $\mathfrak{K} = 1.3$. As the strike increases from at-the-money, less and less mass is left in place (removing mass first from the right tail) and the residual mass of μ is mapped to ν via two functions: one maps the residual mass to the left tail of ν, the other maps the residual mass to the right tail of ν. These functions appear to be increasing for large strikes (Figure 12(b)), but since the transport maps are decreasing for the at-the-money strike (Figure 8(b)), for strikes close to the money these maps could be decreasing 11(b). For strikes lower than the money a mirror-image transport plan emerges where less and less mass stays in place (removing mass first from the left tail) and again the residual mass of μ is mapped to ν via two functions (for brevity we omit the plots).

7. Summary and Conclusion

In this article, we endeavoured to provide a quantitative answer to the question as to whether forward-start options could be effectively replicated using Vanilla products. The take-away message here is that they should rather be thought of as fundamental building blocks, and that trying to replicate them using European options is not reasonable. Our approach using infinite linear programming arguments, makes the methodology directly amenable to computation and calibration to market data; we in particular propose a discretisation scheme to reduce its (infinite) dimensionality, and therefore its complexity. Alternatively, in line with the current active research on robust finance, one could rephrase the problem—using duality arguments—into an optimisation over sets of (martingale) measures, consistent with market data. Several results already exist in that direction, and we leave this for further study.

References

[1] M. Avellaneda. Calibrating volatility surfaces via relative-entropy minimisation. *Applied Math Finance*, 4(1): 37-64, 1997.

[2] D. Baker. Martingales with specified marginals. PhD thesis, Paris 6, tel.archives-ouvertes.fr/tel-00760040v2/document, 2012.

[3] M. Beiglböck, P. Henry-Labordère and F. Penkner. Model-independent bounds for option prices: a mass transport approach. *Finance and Stochastics*, 17(3): 477-501, 2013.

[4] J. Borwein, R. Choksi and P. Maréchal. Probability distributions of assets inferred from option prices via the principle of maximum entropy. *SIAM Journal of Optimization*, 14(2): 464-478, 2003.

[5] J. Borwein, A. S. Lewis. Duality relationships for entropy-like minimization problems. *SIAM Journal of Control and Optimization*, 29(2): 325-338, 1991.

[6] H. Bühler. Applying stochastic volatility models for pricing and hedging derivatives. Presentation available at www.quantitative-research.de/dl/021118SV.pdf, 2002.

[7] L. Campi, I. Laachir and C. Martini. Change of numeraire in the two-marginals martingale transport problem. Preprint, arXiv:1406.6951, 2014.

[8] M.H.A. Davis and D. Hobson. The range of traded option prices. *Mathematical Finance*, 17: 1-14, 2007.

[9] P. Henry-Labordère and N. Touzi. An explicit martingale version of the one-dimensional Brenier theorem. *Finance and Stochastics*, 20(3): 635-668, 2016.

[10] P. Henry-Labordère. Automated Option Pricing: Numerical Methods. *Intern. Journal Theor. App. Fin.*, 16(8): 135-162, 2013.

[11] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2): 327-342, 1993.

[12] D. Hobson. Robust hedging of the lookback option. *Finance and Stochastics*, 2: 329-347, 1998.

[13] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. *Paris-Princeton Lectures on Mathematical Finance 2010*, 2003: 267-318. Springer, Berlin, 2010.

[14] D. Hobson and M. Klimmek. Robust price bounds for the forward starting straddle. *Finance and Stochastics*, 19(1): 189-214, 2015.

[15] D. Hobson and A. Neuberger. Robust bounds for forward start options. *Mathematical Finance*, 22(1): 31-56, 2012.

[16] D.G. Luenberger and Y. Ye. Linear and Nonlinear Programming. Springer US, 2008.

[17] S. Mehrotra and D. Papp. Generating nested quadrature formulas for general weight functions with known moments. Preprint, arXiv:1203.1554, 2012.

[18] J. Obłój. The Skorokhod embedding problem and its offspring. *Probability Surveys*, 1: 321-390, 2004.

[19] A. Shapiro. Semi-infinite programming, duality, discretisation and optimality conditions: Invited Review. *Optimization*, 58(2): 133-161, 2009.

[20] S. Shreve. Stochastic calculus for finance I. The binomial asset pricing model. Springer, 2004.

[21] G. Still. Discretization in semi-infinite programming: the rate of convergence. *Mathematical Programming*, 91(1): 53-69, 2001.

[22] V. Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36: 423-439, 1965.

[23] K. Tanaka and A. Toda. Discrete approximations of continuous distributions by maximum entropy. *Economics Letters*, 118: 445-450, 2013.

Department of Mathematics, Imperial College London

E-mail address: sergey.badikov08@imperial.ac.uk, a.jacquier@imperial.ac.uk, q.liu12@imperial.ac.uk, p.roome11@imperial.ac.uk